

# Carnegie Mellon University HeinzCollege

INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

**Discussion of:**

**Big Data Based Government Economic  
Monitoring and Targeted Action**

**Amelia M. Haviland**  
Professor of Statistics and Public Policy



# ***Big Data Analytics + AI → Social Good***

- AI & Big Data Analytics so far have mostly been put to use by those at the top of the economic ladder and for their gain
- Putting them to use for Social Good is still a wide-open space
  - Government Budget Planning for Social Services
  - Government Accountability, Performance Assessment
  - Transforming the process to be more nimble, responsive, adaptable
  - Targeting service provision to be more effective

# ***Big Ideas, New Ideas***

- Many data sources, most not designed for these purposes
  - Focus less on each data source, more on what the set can provide
  - In contrast to primary data collection – cons: time lags, low response rates, pros: known and usually strong measurement properties
- Block Chain related technologies for Privacy
  - I have not seen this technology applied in this way in this space before

# Opportunities & Challenges

- Looking at this process in new ways has great potential
- Also can result in suggestions that are not improvements or can create new (sometimes worse) problems.
  - ML and HealthCare
  - Facial recognition software & Parole
- **Some questions to ask:**
  - Does this provide them new information?
  - Is this information actionable, should it be acted on?
  - How much would the organization have to change to make it actionable?
  - Is that change wanted/feasible?

# Group Risk versus Individual at Risk

- Predicting Flu Rates over time – Huge Success, Easier Problem
  - How many in this area have/will have the flu?
  - Who in this area has the flu?
- Hierarchy of Risk for Social Problems
  - *Algorithmic Risk = Different*
  - Population at Risk due to known precursors to the social problem
  - Population in Need – experiencing the social problem
  - Population in Demand – those need who are willing to accept services

# Dimensions of Privacy

- Direct versus Indirect Identification
  - Direct: Name, Social Security Number – unique identifiers
  - Indirect: date of birth, location, health status
- Harms from Identification
  - *Legal system – illegal drugs, domestic abuse – these are crimes*
  - *Stigma about the social problem, stigma about being identified as needing help*
  - *Public exposure, financial or job risk*